Claims

What is claimed is:

1   1.   A method of assembling a life science knowledge base comprising the steps of:

2   (a)   generating two or more nodes indicative of life science data using

3   a life science taxonomy;

4   (b)   assigning to one or more pairs of nodes a representation descriptor,

5   the representation descriptor corresponding to a relationship between a pair of

6   nodes;

7   (c)   assembling two or more nodes and one or more representation

8   descriptors assigned to one or more pairs of said two or more nodes into an

9   electronic database such that at least one of said two or more nodes is joined to a

10   another node by a representation descriptor.

1   2.   The method of claim 1 further comprising the step of receiving life science data,

2   wherein the step of generating two or more nodes is based at least in part on said

3   received life science data.

1   3.   The method of claim 2, wherein the step of receiving life science data comprises

2   collecting said life science data using a software agent.

1   4.   The method of claim 2, wherein the step of receiving life science data further

2   comprises receiving one or more of metadata and context data.

1   5.   The method of claim 1, wherein said life science data comprises information

2   representative of a molecule, biological structure, physiological condition, trait,

3   phenotype, biological process, clinical data, medical data, or disease data and

4   chemistry.

1 6.   The method of claim 1, wherein said life science data comprises a descriptor of

2   the condition, location, amount, or substructure of a molecule, biological

3   structure, physiological condition, trait, phenotype, biological process, clinical

4   data, medical data, or disease data and chemistry.

1 7.   The method of claim 1, wherein the step of generating two or more nodes

2   comprises reformatting at least a portion of said life science data.

1 8.   The method of claim 1, wherein one or more of the representation descriptors

2   correspond to a epistemological relationship between a pair of nodes.

1 9.   The method of claim 1, wherein one or more of the representation descriptors

2   comprise a case frame.

1 10.   The method of claim 1 further comprising the step of providing an ontology for

2   use with representation descriptors, wherein the step of assigning to one or more

3   pairs of nodes a representation descriptor is based on said ontology.

1 11.   The method of claim 1, further comprising the step of segregating said electronic

2   database into two or more sectors such that access may be restricted to one or

3   more selected sectors.

1 12.   The method of claim 1, wherein at least one of the two or more nodes itself

2   represents a representation descriptor.

51

1 13. An article of manufacture having a computer-readable program carrier with

2   computer-readable instructions embodied thereon for performing the method of

3   claim 1.

1 14. A system for assembling a life science knowledge base comprising:

2   (a)   a data collector configured to receive life science data and to

3   generate nodes based on said life science data;

4   (b)   a relationship generator configured to assign a relationship

5   descriptor to a pair of nodes, the representation descriptor corresponding to a

6   relationship between a pair of nodes; and

7   (c)   a knowledge assembler configured to assemble two or more nodes

8   and one or more representation descriptors assigned to one or more pairs of said

9   two or more nodes into an electronic database such that at least one of said two or

10   more nodes is joined to another node by a representation descriptor.

1 15. The system of claim 14 further comprising a graphical user interface configured

2   to permit a user to query the electronic database at least on the relationship

3   between at least two nodes.

1 16. The system of claim 14 further comprising a data input interface configured to

2   permit a user to submit life science data to the data collector.

1 17. The system of claim 16, wherein the data input interface is further configured to

2   permit a user to assign a representation descriptor to a pair of nodes in the

3   electronic database.

1   18.   The system of claim 14 further comprising an access manager configured to

2   restrict access of a user to one or more portions of the electronic database.

1   19.   The system of claim 14 further comprising a software agent in electronic

2   communication with the data collector, wherein the software agent is configured

3   to collect life science data.

1   20.   The system of claim 14, wherein said life science data comprises information

2   representative of a molecule, biological structure, physiological condition, trait,

3   phenotype, biological process, clinical data, medical data, or disease data and

4   chemistry.

1   21.   The system of claim 14, wherein said life science data comprises a descriptor of

2   the condition, location, amount, or substructure of a molecule, biological

3   structure, physiological condition, trait, phenotype, biological process, clinical

4   data, medical data, or disease data and chemistry.

1   22.   The system of claim 14 further comprising a library of machine readable

2   representation descriptors in electronic communication with the relationship

3   generator.

1   23.   The system of claim 14, wherein one or more of the representation descriptors

2   correspond to a epistemological relationship between a pair of nodes.

1   24.   The system of claim 14, wherein one or more of the representation descriptors

2   comprise a case frame.

1   25.   The system of claim 14, wherein one or more of the nodes represents a

2   representation desctiptor.

1   26.   A computer program product comprising:

2   an electronic database storing a plurality of case statements, each case

3   statement comprising;

4   an first object identifier;

5   a relationship connector; and

6   a second object identifier

7   wherein the relationship connector is based on a life science ontology.

1   27.   The product of claim 26, wherein a set of said case statements define a biological

2   function.

1   28.   The product of claim 27, wherein the biological function comprises a chemical

2   reaction.

1   29.   The product of claim 27, wherein the biological function comprises transport.

1   30.   The product of claim 27, wherein the biological function comprises digestion of a

2   biomolecule.

1   31.   The product of claim 26, wherein at least one of the first and second object

2   identifiers identifies a biomolecule.

1   32.   The product of claim 26, wherein at least one of the first and second object

2   identifiers identifies a biological function.

1 33. The product of claim 26, wherein at least one of the first and second object

2 identifiers identifies a relationship connector.

1 34. The product of claim 26, wherein a the relationship connector represents an

2 identity relationship.

1 35. The product of claim 26, wherein a the relationship connector represents a

2 product relationship.

1 36. The product of claim 26, wherein a the relationship connector represents a

2 substrate relationship.

1 37. The product of claim 26, wherein a the relationship connector represents a

2 enzymatic relationship.

1 38. The product of claim 26 further comprising a graphical user interface configured

2 to permit a user to query the database at least on the relationship between

3 biological object identifiers.

1 39. The product of claim 26 further comprising a data input interface configured to

2 permit a user to create case statements.

1 40. The product of claim 26 further comprising an access manager configured to

2 restrict access of a user to one or more portions of the electronic database.

1 41. A method of providing knowledge about and permitting investigation of a

2 functional physicochemical life science pathway comprising the steps of:

3   (a)   generating plural nodes indicative of life science data using a life

4   science taxonomy;

5   (b)   assigning to respective pairs of nodes case frames corresponding to

6   a physical or functional relationship between the node pairs;

7   (c)   assembling the nodes and case frames to produce an addressable

8   electronic life science knowledge base representing a web of nodes interrelated by

9   case frames; and

10   (d)   providing a user interface including a knowledge base search

11   engine to permit a user to locate and obtain data descriptive of a selected region of

12   said web.

1   42.   A method of converting life science data into life science knowledge comprising

2   the steps of:

3   (a)   providing an electronically accessible knowledgebase

4   representative of a web of life science information comprising a plurality of nodes

5   indicative of life science data interrelated by a plurality of case frames

6   representative of the functional or physical relationship between pairs of nodes;

7   and

8   (b)   providing through an electronic link to a plurality of life science

9   data providers a data input interface which permits diverse providers to input new

10   data into said knowledgebase, requires input of the new data to fit within one or

11   more of a plurality of predetermined node categories, and requires input of the

12   new data to fit within one or more of a plurality of case frames, thereby to enforce

13 a consistent data structure and ontology rules and to permit iterative accumulation

14 of new data linked functionally within said knowledgebase.

1 43. A method for representing life science knowledge, the method comprising using a

2 case frame to represent the life science knowledge, the case frame comprising:

3   a first object identifier;

4   a relationship connector; and

5   a second object identifier,

6 wherein the relationship connector is based on a life science ontology.

1 44. The method of claim 43, wherein the case frame represents a biochemical

2 reaction.

1 45. The method of claim 44, wherein the biochemical reaction is an enzymatic

2 reaction.

1 46. The method of claim 43, wherein the case frame represents binding.

1 47. The method of claim 46, wherein the binding comprises nucleic acid binding.

1 48. The method of claim 43, wherein the case frame represents an agglomeration of

2 molecules.

1 49. The method of claim 43, wherein the case frame represents a modification of a

2 polymer.

1 50. The method of claim 49, wherein the modification of a polymer comprises a post-

2 translational modification of a protein.

1   51.   The method of claim 50, wherein the post-translational modification is selected

2   from a group consisting of phosphorylation, acetylation, peptide-bond cleavage,

3   glycosylation, lipidation, fatty-acylation, prenylation, methylation, metallation,

4   cross-linking, hydroxylation, sulfation, ADP-ribosylation, and covalent

5   attachment of prosthetic groups.

1   52.   The method of claim 43, wherein the case frame represents competition between

2   processes.

1   53.   The method of claim 43, wherein the case frame represents requirements of a

2   process.

1   54.   The method of claim 43, wherein the case frame represents translocation.

1   55.   The method of claim 43, wherein the case frame represents activation.

1   56.   The method of claim 55, wherein the case frame represents direct activation.

1   57.   The method of claim 43, wherein the case frame represents inhibition.

1   58.   The method of claim 57, wherein the case frame represents direct inhibition.

1   59.   The method of claim 43, wherein the case frame represents patient data.

1   60.   The method of claim 59, wherein the case frame represents patient data of a

2   phynotypic nature.

1   61.   The method of claim 59, wherein the case frame represents patient data of a

2   genotypic nature.

1 62. The method of claim 43, wherein the case frame represents a product relationship.

1 63. The method of claim 43, wherein the case frame represents an experiment.

1 64. The method of claim 63, wherein the relationship connector identifies linkages to

2 characteristics and parameters of the experiment.

1 65. The method of claim 43, wherein the case frame represents an animal model.

1 66. The method of claim 43, wherein the case frame represents a tissue type.

1 67. The method of claim 43, wherein the case frame represents a SNP.

1 68. The method of claim 43, wherein the case frame represents a splice varient.

1 69. The method of claim 43, wherein the case frame represents microRNA.

1 70. The method of claim 69, wherein the relationship connector identifies linkages to

2 a function of the microRNA.

1 71. The method of claim 43, wherein the first object identifier represents a

2 chromosome.

1 72. The method of claim 43, wherein the case frame represents a chromosomal

2 modification.

1 73. The method of claim 43, wherein the case frame represents a silenced gene.

1 74. The method of claim 43, wherein the case frame represents a subcellular

2 compartment.

1   75.   The method of claim 43, wherein the relationship connector represents a subset

2   relationship.

1   76.   The method of claim 43, wherein the relationship connector represents a

2   positional relationship.

1   77.   The method of claim 43, wherein the relationship connector identifies a drug

2   combination for treatment of one or more conditions.

1   78.   The method of claim 43, wherein the case frame represents a disease.

1   79.   The method of claim 78, wherein the disease is cancer.

1   80.   The method of claim 43, wherein the case frame represents a population.

1   81.   The method of claim 43, wherein the case frame represents a xenograft model.

1   82.   The method of claim 43, wherein the case frame represents a cell line.

1   83.   The method of claim 43, wherein the relationship connector represents that an

2   entity is a biomarker.

1   84.   The method of claim 43, wherein the first object identifier identifies a protein.

1   85.   The method of claim 43, wherein the first object identifier identifies a gene.

1   86.   The method of claim 43, wherein the first object identifier identifies a metabolite.

1   87.   The method of claim 43, wherein the first object identifier identifies a population.

2   88.   The method of claim 43, wherein the first object identifier identifies an epidemic.

1   89.   The method of claim 43, wherein the first object identifier identifies a pathogen.

1   90.   The method of claim 89, wherein the pathogen is selected from a group consisting

2   of a virus, a bacteria, a fungus, and a prion.

1   91.   The method of claim 89, wherein the relationship connector represents the

2   implications of a specific disease.

3